Electronic Appendix

This is the electronic appendix to the article

Bovine tuberculosis in cattle: reduced risk on wildlife-friendly farms

F. Mathews[a,*], L. Lovett[2], S.J. Rushton[2] & D.W. Macdonald[1]

[1] *Wildlife Conservation Research Unit, University of Oxford, Tubney House, Abingdon Road, Tubney, OX13 5QL, UK.*

[2] *Centre for Life Sciences Modelling, University of Newcastle-upon-Tyne, Devonshire Building, Newcastle-upon-Tyne, NE1 7RU, UK.*

Author for correspondence (Fiona.mathews@zoo.ox.ac.uk)

MATERIALS AND METHODS

Data

Diagnosis of bTB in the UK is conducted using a combination of the tuberculin skin test and microbiological culture. Herd breakdowns are considered 'confirmed' by the Department for Environment, Food and Rural Affairs (DEFRA) if *Mycobacterium bovis* is cultured from at least one animal *post mortem*. Because of the imperfect sensitivity of the diagnostic methods, a small proportion of negative reactions to the skin test will be false negatives. Breakdowns are confirmed by culture in approximately two-thirds of cases with positive skin tests. Some of the discrepancy is due to under-diagnosis by culture, and some is due to false-positive results from the skin test. For this study, 'case' farms were those with one or more confirmed breakdowns since 1997 in addition to at least two instances of positive reactors to the skin test since 1994. Farms with breakdowns caused by the importation of infected cattle, as defined by DEFRA, were excluded. Control farms had no breakdowns (confirmed or unconfirmed) since 1994. The registered grid references for each case herd (as opposed to the mailing address of the farm) were plotted using a geographical information system (GIS; Arcview 3.2, Environmental Systems Research Institute). The two 1000 sq km areas with the highest density of case farms were selected for further study: these were in North Devon and the Herefordshire/Gloucestershire border. Thirty case and 30 control farms were randomly selected from all those present within each study area.

To permit the analysis of risk of repeated bTB breakdown events – something not possible with the original dataset because of there having been only 9 repeat breakdown events – additional data on the study farms were obtained for the years 2000-2004 inclusive. Farms with repeated breakdown events were defined as those with more than 1 confirmed breakdown since 1997. Control farms, as previously, were defined as those with no cattle reacting to the tuberculin skin test since 1994. There were 38 farms with repeated breakdowns and 41 controls.

Hedgerow characteristics were obtained from aerial photographs with a resolution of 1m taken in the year 2000 (Get Mapping®UK, Witney, UK) which were examined using the GIS package Geographic Resources Analysis Support System (GRASS) (Westervelt 1990). The total length of hedgerow per farm was measured, and twenty hedgerows were examined in more detail. These hedgerows were selected on the basis of their proximity to the nodes of a 250m grid. Land cover information was obtained from remotely-sensed satellite data (Fuller *et al*. 1994) that categorise land cover into 25 classes at 25m resolution. Spatial statistics were computed using ARCVIEW 3.2 (McGarigal & Marks 1994). Shannon's diversity index was used to describe the diversity of patches within each farm (Magurran 1988). The index equals zero when there is only one patch. The distance to the nearest infected farm (defined as for case farms) was measured, using all case farms within the study area, not just the 30 included in the analysis.

Summary variables for landcover and boundary variables were computed using Principal Components Analysis to overcome the unit-sum constraint that would restrict analysis of bTB breakdown response to individual habitat features (PCA).

Data on farm extent and land use were available for 39 case and 19 control farms from the Veterinary Laboratories Agency's database. This validated the reports of farm size given to the Rural Payments Agency (Pearson's correlation coefficient 0.966, *P*<0.001), and indicated that most (87%, *SD* 17%) of the land area on each holding was used for cattle.

Model selection using Akaikes Information Criterion

The main objective in regression modelling is to seek to explain the maximum variation in the response variable (here herd breakdowns) with the minimum number of explanatory variables (habitat characteristics) under the implicit assumption that the predictor variables are key determinants of the process being investigated. Traditionally, model selection is based on forward or backward selection of predictor variables with inclusion in the model determined by a significance test based on a cut off level (often *P*<0.05). The choice of the *P*-value is arbitrary, its value depends on sample size, and its magnitude does not provide direct evidence of the validity of the alternative hypothesis, but rather the consistency of the data with the null hypothesis (Ellison 1996). Information-theoretic approaches differ in that they assume that the modeller has identified the suite of predictors a priori and as such the best models are identified by searching the suite of models that can be constructed from the predictors. The approach involves comparing different models by assessing their **relative** differences to the unknown true mechanism, using an information statistic termed the Akaike Information Criterion (AIC).

$$AIC_c = -2\ (\text{log likelihood}) + 2K$$

AIC combines the maximum log-likelihood of the model, which is an estimate of the probability of observing the real data given the model (and a measure of lack of model fit) with, K, the number of parameters used in it. The value of the criterion for any model is not of itself important. Rather the relative difference in AICs between competing models allows the models to be ranked. Thus, where two competing models have different log-likelihoods and the same number of predictors, the lowest will be ranked best and indeed is the model which explains most variation in the observed response (herd breakdowns). In contrast with traditional methods, inclusion of the number of parameters in the estimate of AIC, however, enforces a principle of parsimony in the modelling since it effectively imposes a penalty for including variables.

If two models have similar AIC values then there is very little evidence to suggest that they are any different in their ranking. Here we calculated AIC_c (which is AIC adjusted for small sample sizes (Hurvich & Tsai 1989)) for each model and used the methodology of Burnham and Anderson (2002) to assess the set of models created. First, we compared the difference between the AIC_c of each model and the one with the lowest AIC_c in the candidate set to provide the ranking metric (Δ AIC_c). There is increasing evidence for the rejection of lower ranking models as the Δ AIC_c increases . (Burnham & Anderson 2002). Models with Δ AIC_c values <2 were selected as the most parsimonious (Anderson *et al.* 2000, 2001, Burnham & Anderson 2002). Second, Akaike weights (w) and evidence ratios (w_j/w_i) were derived to give an assessment of how all of the models compared to the top ranking model. The Akaike weights can be interpreted, heuristically, as the probability of the candidate model being the 'best' out of all those considered, given the data. The evidence ratios therefore show how much 'better' the top-ranking model is than the candidate model.

$$w=\exp(-\Delta_i/2)/\sum_{R}^{r=1}\exp(-\Delta_r/2)$$ (Anderson *et al*. 2000)

The relative importance of each individual predictor variable in the final suite of models was then assessed by calculating predictor weights for each by summing the *w* of each model in which a predictor appeared.

Standardised residuals of each model were examined to ensure conformity to the model assumptions. Model fit was assessed using the model χ^2. All of the models where the difference in AICc to the best model was less than, all had acceptable goodness-of-fit (worst fit was for model bTB=hedgepc2, χ^2=5.85, df=1, P=0.016).

Results

Recurrence of bTB

Recurrence of bTB breakdowns on farms previously cleared of the disease is an increasing problem, and it is unclear whether the aetiology is the same as for single breakdowns. We therefore analysed whether the same predictors could explain variations in the risk of repeated bTB events. For this, additional data for the study farms was obtained from the years 2000-2004 inclusive. The results were similar to those previously described: herd size, the proximity of the next nearest farm with bTB, hedgerow score, and the number of reported road-killed badgers were again the key parameters. The adjusted odds ratios from multiple logistic regression of the top-ranking model were: herdsize 1.01, P=0.006; nearcase 0.67, P=0.090, hedgepc2 2.20, P=0.024 badgers 1.51, P=0.051; 73.4% of cases correctly classified; R^2=0.435. Again the density of cattle within a farm, and factors relating to topography, land cover, and the type and extent of pasture and woodland were of little or no predictive value.

Discussion

Habitat diversity

There is some suggestion that an increase in habitat diversity (as measured by Shannon's Diversity Index) is positively associated with bTB risk, but the evidence is weak. Previous research has found inconsistent relationships between habitat heterogeneity and the incidence of repeated badger control operations, with the direction of the association differing between land classes (White *et al.* 1993).

The relationships between habitat factors, cattle grazing and badger excretory behaviour

Although cattle generally avoid grazing areas contaminated by badger faeces and urine (Benham & Broom 1989), this avoidance is balanced by a strong preference to graze herbaceous plants and longer grass (Hutchings & Harris 1997). Thus when there is little intra-herd competition for long forage – for example when cows are first turned onto long pasture - there is strong avoidance of active latrines. However, this avoidance diminishes, or even disappears altogether when less long grass is available. The lush grass around latrines and urination sites is then readily consumed (Hutchings & Harris 1997). Good hedgerows provide a rich source of long forage (indeed, these areas are grazed out first in preference to the centre of the field (Hutchings & Harris 1997)). It therefore follows that the greater the availability of good hedgerow, the lower the risk of bTB transmission due to cattle grazing areas of contaminated pasture. The density of pasture edge (which generally, like hedgerows, provides areas of longer grass) was also negatively correlated with bTB risk, lending some support to this hypothesis. If swards remain longer on hedgerow-rich farms then the pasture will also be less attractive to badgers as foraging habitat (Kruuk *et al*. 1979) and less contaminated with single deposits of excreta away from latrines. This could be important since cattle appear not to avoid this type of contamination (Hutchings & Harris, 1997).

There is debate about the extent to which cattle avoid pasture contaminated with urine only (such as the scent-marked areas where badgers cross boundary features) as opposed to latrines which also contain faeces (Benham & Broom 1989; Hutchings & Harris 1997). Indeed it has been suggested that greater linear feature density increases the absolute number of urinations on pasture (at the points where badgers cross the boundary features), thereby increasing the risk of bTB transmission to cattle (White *et al*. 1993). However, when the data from this study are re-analysed without a single extreme outlier, no relationship is apparent (r=0.15, P=0.389,df=1, 35). On the contrary, we suggest that increased hedgerow length may actually reduce contact rates between cattle and infected excretory products from badgers.

Table 1. Component matrix for summary hedgerow and landcover variables[a] derived using principal components analysis.

Variable	Component 1			Component 2		
	Standardized factor score coefficients	Eigen value	% variance explained	Standardized factor score coefficients	Eigen value	% variance explained
Total hedgerow length (km) $100ha^{-1}$	0.552			-0.352		
Mean hedgerow length (m)[b]	0.116			0.878		
Mean score for connections $100m^{-1}$ (score 1 for each connection to hedgerow and 2 for each connection to woodland)[b]	0.591			0.157		
		1.496	49.9		1.087	36.2
Water	0.000			0.000		
Grass heath	-0.003			0.004		
Grazed/mown turf	-0.218			0.785		
Semi natural grassland/meadow	0.429			0.057		
Rough/marsh grasss	0.000			-0.002		
Bracken	0.004			0.000		
Dense shrub heath	0.000			-0.001		
Scrub/orchard	-0.001			0.001		
Deciduous woodland	0.211			-0.154		

Coniferous woodland	-0.001			-0.001		
Tilled land	-0.414			-0.549		
Suburban/rural development	-0.010			-0.018		
Urban development	0.000			0.000		
Felled woodland	-0.001			-0.001		
Open shrub heath	-0.001			-0.001		
		3.4^{10}	55.1		1.6^{10}	26.1

[b]The fifteen landcover types that occurred in the study areas (out of the possible 25) were included in the anaysis.

[a]Hedgerow lengths and number of end connections were means computed for 20 hedgerows.

Table 2. Descriptive data for candidate habitat variables considered in models.

Variable	Description	Mean (SD)
Hedgepc1	Principal component 1 for hedgerow availability (see Table 1)	0 (1)
Hedgepc2	Principal component 2 for hedgerow availability (see Table 1).	0 (1)
Width	Modal width (m) per hedge derived from measurements taken every 100m. Mean computed across 20 hedgerows.	4.2 (2.8)
Gaps	Number of gaps per 100m of hedgerow (mean for 20 hedgerows)	0.5 (0.4)
Standards	Number of standard trees i.e. those emerging above level of hedgerow, per 100m (mean for 20 hedgerows)	2.4 (1.2)
Head	Number of conservation buffer strips (headlands) at least 2m wide adjacent to hedgerow from which cattle are excluded. Can be 0, 1 or 2 per hedgerow (mean for 20 hedgerows).	0.1 (0.1)
Coverpc1	Principal component 1 for landcover (see Table 1)	0 (1)
Coverpc2	Principal component 2 for landcover (see Table 2).	0 (1)
Domcov	Dominant land cover type	-
Decid	Amount of deciduous woodland (land cover class 16)(ha)	9.8 (9.8)
Decidedge	Length of edge of deciduous woodland (km)	5.2 (4.8)
Hetdecid	Heterogeneity of woodland (perimiter (m) :area (ha) ratio)	0.4 (0.2)
Turf	Area of mown or grazed grassland (land cover class 6) (ha)	25.8 (12.1)
Turfedge	Length of edge of mown or grazed grassland (km)	10.3 (3.4)
Hetturf	Heterogeneity of mown or grazed grassland (perimiter: area ratio)	0.4 (0.1)
Grass	Total area of grassland excluding rough grass (sum of land cover classes 6 and 7 ie mown/grazed turf &	54.0 (13.5)

	meadow/verge/semi-natural grassland). Summary variable corresponds to 'Key cover type' (Fuller *et al*. 1994).	
Grassedge	Length of edge of all grass (km)	25.0 (5.5)
Hetgrass	Heterogeneity of all grass (perimiter: area ratio)	0.5 (0.1)
LPI	Largest patch index (% of total area)	25.4 (11.6)
NumP	Number of habitat patches	97.9 (24.0)
MPS	Mean habitat patch size (ha)	1.1 (0.3)
Edgedensity	Density of patch edges (length (m)/total farm area (ha))	289.3 (36.8)
SDI	Shannon's diversity index. A relative measure of patch diversity. The index equals zero when there is only one patch in the landscape and increases as the number of patch types or proportional distribution of patch types increases (McGarigal & Marks 1994).	1.7 (0.2)
MSI	Mean shape index. A metric describing the shape complexity of habitat patches. MSI is greater than one, MSI = 1 when all patches are circular (polygons) or square (grids). MSI = sum of each patches perimeter divided by the square root of patch area (ha) for all patches, and adjusted for circular standard (polygons), or square standard (grids), divided by the number of patches (McGarigal & Marks 1994).	1.4 (0.1)
County	Two geographical areas: Devon or Hereford/Shropshire	-
Maxalt	Maximum altitude on farm (m). Derived from contour maps.	132.0 (66.3)
Minalt	Minimum altitude on farm. Derived from contour maps	89.0 (55.4)
Altdiff	Difference between maxalt and minalt	43.0 (25.2)
Herdsize	Number of cattle registered	201.6 (187.1)

Area	Size of farm as registered with DEFRA's Rural payment's agency	108.5 (92.9)
Herdensity	Number of cattle/area of holding.	2.9 (6.6)
Nearcase	Distance (km) to nearest 'case' herd from herd's registered grid co-ordinates.	2.6 (1.3)
Badgers	Number of road traffic accident records for badgers within the 1km square containing the herd's registered grid co-ordinates. The number of records to some extent reflects the local community's efforts at reporting and recording casualties, and this is associated with local history of bovine TB in cattle.	1.3 (2.0)
Badger5k	Number of road traffic accident records for badgers within the 5km square containing the herd's registered grid co-ordinates. The number of records to some extent reflects the local community's efforts at reporting and recording casualties, and this is associated with local history of bovine TB in cattle.	32.9 (38.1)
Badgerpres	Presence/absence of badger road traffic accident records for badgers within the 1km square containing the herd's registered grid co-ordinates.	-
Badgerpres5	Presence/absence of badger road traffic accident records for badgers within the 5km square containing the herd's registered grid co-ordinates.	

Spatial statistics were computed in the Geographical Information System Arcview 3.2 using the Frag-Stats interface (McGarigal & Marks 1994).

Table 3. Akaike Information Statistics for logistic regression models relating bTB incidence in cattle herds to habitat predictors alone. Models are ranked from the most plausible (Δ AIC_c=0) to least plausible. All models with Δ AIC_c <2.05 are listed. W_i/w_j indicates the likelihood of the top-ranking model compared with the model on a given row. The overall % correct classification ranges from 58.3 to 66.7 (mean 65.2% correct presence and 60.8% correct absence).

Model	AIC_c[a]	Δ AIC_c	W[b]	w_i/w_j[c]	R^2[d]
Hedgepc2, head, gaps[e]	162.61	0.00	0.113	1.0	0.13
Hedgepc2, width, head, gaps	163.16	0.55	0.086	1.31	0.14
Hedgepc2, width, gaps	163.47	0.86	0.074	1.53	0.12
Hedgepc2, head, gaps, SDI	163.81	1.20	0.062	1.82	0.14
Hedgepc2, head	164.05	1.44	0.055	2.05	0.09
Hedgepc2, head, gaps, turfedge	164.15	1.54	0.052	2.20	0.13
Head, gaps	164.15	1.54	0.052	2.16	0.09
Hedgepc2, gaps	164.17	1.56	0.052	2.18	0.09
Hedgepc2, head, gaps, hetturf	164.26	1.65	0.050	2.28	0.13
Hedgepc2, head, gaps, width, SDI	164.28	1.66	0.049	2.30	0.16
Hedgepc2, width, gaps, SDI	164.33	1.72	0.048	2.37	0.13
Hedgepc2, head, gaps, coverpc1	164.38	1.77	0.047	2.42	0.13
Gaps	164.38	1.77	0.047	2.42	0.04
Hedgepc2, head, gaps, decidedge	164.50	1.89	0.044	2.57	0.13
Hedgepc2, gaps, head, hetwood	164.50	1.89	0.044	2.57	0.13
Width, head, gaps	164.54	1.93	0.043	2.62	0.11
Hedgepc2, head, gaps, LPI	164.58	1.96	0.042	2.67	0.13
Hedgepc2	164.61	2.00	0.042	2.71	0.06

[a] Akaike's Information Criterion adjusted for small sample sizes.

[b] Akaike weight.
[c] w_i/w_j, evidence ratio.
[d] Nagelkerke's R-square (Nagelkerke 1991).

[e] See Table 1 for variable descriptions.

Anderson, D.R., Burnham, K. , & Thompson, W. L. 2000 Null hypothesis testing: problems, prevalence and an alternative. *J. Wild. Manag*. **64**, 912-923.

Anderson, D.R., Burnham, W. R., Gould, W.R. & Cherry, S. 2001 Concerns about finding effects that are actually spurious. *Wild. Soc. Bull*. **29**, 311-316.

Benham, B. F. J. & Broom, D. M. 1989 Interactions between cattle and badgers at pasture with reference to bovine tuberculosis transmission. *Br. Vet. J.* **145**, 226-241.

Burnham, K. P. & Anderson, D. R. 2002 *Model Selection and Multimodel Inference: a practical information-theoretic approach, 2nd edition.* (Springer-Verlag, New York).

Ellison, A. M. 1996 An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecol. Appl*. **6**, 1036-1046.

Fuller, R.M., Hall, J.R.B., Groom, G.B. & Parr, T.W. 1994 The Land Cover map of Great Britain and its applications. *Annual Report of the Institute of Terrestrial Ecology*. (NERC, Swindon, UK,).

Hurvich, C. M. & Tsai, C. 1989 Regression and time series model selection in small samples. *Biometrika* **76**, 297-307.

Hutchings, M. R. & Harris, S. 1997 Effects of farm management practices on cattle grazing behaviour and the potential for transmission of bovine tuberculosis from badgers to cattle. *Vet. J*. **153**, 149-162.

Kruuk, H., Parish, T., Brown, C.A.J. & Carera, J. 1979 The use of pasture by the European badger (*Meles meles). J. Appl. Ecol*. **16**, 453-459.

Magurran, A. E. 1988 *Ecological diversity and its measurement*. (Chapman & Hall, London).

McGarigal, K. & Marks, B. 1994 *FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure*. General Technical Report PNW GTR-351. (United States Department of Agriculture Forest Service, USA).

Nagelkerke, N.J.D. 1991 A note on a general definition of the coefficient of determination. *Biometrika* **78**, 691-692.

Westervelt, J.M., Shapiro, M., Goran, W. & Gerdes, D. 1990 *Geographic Resource Analysis Support System. Version 4.0 User's Reference manual.* (USACERL ADP Report N-87/22, USA).

White P. C. L., Brown, J. A. & Harris, S. 1993 Badgers (*Meles meles*), cattle and bovine tuberculosis (*Mycobacterium bovis*): a hypothesis to explain the influence of habitat on the risk of disease transmission in southwest England. *Proc. R. Soc. Lond. B.* **235**, 277-284.